Global temperature change and its uncertainties since 1861

C. K. Folland[1], N. A. Rayner[1], S. J. Brown[1,2], T. M. Smith[3], S. S. P. Shen[4], D. E. Parker[1], I. Macadam[1], P. D. Jones[5], R. N. Jones[6], N. Nicholls[7] and D. M. H. Sexton[1]

Abstract. We present the first analysis of global and hemispheric surface warming trends that attempts to quantify the major sources of uncertainty. We calculate global and hemispheric annual temperature anomalies by combining land surface air temperature and sea surface temperature (SST) through an optimal averaging technique. The technique allows estimation of uncertainties in the annual anomalies resulting from data gaps and random errors. We add independent uncertainties due to urbanisation, changing land-based observing practices and SST bias corrections. We test the accuracy of the SST bias corrections, which represent the largest source of uncertainty in the data, through a suite of climate model simulations. These indicate that the corrections are likely to be fairly accurate on an annual average and on large space scales. Allowing for serial correlation and annual uncertainties, the best linear fit to annual global surface temperature gives an increase of 0.61 ± 0.16°C between 1861 and 2000.

Estimating Uncertainties in Temperature Data

Land surface-air temperature (LAT) and SST observations are taken from a new global data set (HadCRUTv, *Jones et al.*, 2001) whose variance is homogenised for local temporal variations in data density. To assess the uncertainties in annual global and hemispheric average surface temperature anomalies due to data gaps, random data representivity errors and measurement errors, we employ a two-step optimal averaging (OA) method. The OA method provides a better estimate of the true mean than does a simple average and a consistent framework on which to add independent uncertainties due to other factors described below.

Step 1 uses standard correlation functions to estimate the local covariability of the data (*Smith et al.*, 1994). It then uses annual 5° latitude by 5° longitude grid-box HadCRUTv anomalies (relative to the 1961-1990 average) and their uncertainties, to create averages and uncertainties on a coarser 10° latitude by 20° longitude grid. Step 2, "reduced space" OA (RSOA, *Shen et al.*, 1998), takes the coarse-grid averages and error estimates and utilises a weakly truncated set of global or hemispheric Empirical Orthogonal Functions (EOFs) calculated on the coarse grid to calculate global or hemispheric averages and their uncertainties. In both steps, averages were estimated by

$$\hat{\bar{T}} = \sum_{i \in N} w_i T_i$$

with the constraint that the optimal weights w_i sum to unity. Here T_i is the temperature anomaly at grid box i and N represents the network of boxes with observed data. The optimal

[1]Hadley Centre, Met Office, Bracknell, UK
[2]Now at Dept. Environment, Transport & Regions, London, UK
[3]National Climatic Data Center, Asheville, NC
[4]Dept. Mathematical Sciences, Univ. Alberta, Edmonton, Canada
[5]Climatic Research Unit, Univ. East Anglia, Norwich, UK
[6]Commonwealth Sci. and Indust. Res. Org., Aspendale, Australia
[7]Bureau of Meteorology Research Centre, Melbourne, Australia

weights are used to minimize and estimate the mean squared error, $\overline{\varepsilon^2}$, of the data. This has an EOF representation in RSOA given by (*Shen et al*., 1998):

$$\overline{\varepsilon^2} = \left\langle \left(\overline{T} - \hat{\overline{T}} \right)^2 \right\rangle \approx \sum_{n=1}^{M} \lambda_n \left[\overline{\psi_n} - \sum_{i \in N} w_i \psi_n(i) \right]^2 + \sum_{i \in N} w_i^2 \left\langle E_i^2 \right\rangle$$

where $\overline{T}$ is the true average to be estimated, λ_n is the eigenvalue of EOF mode *n*, $\psi_n(i)$ is the value of EOF mode *n* at coarse grid box *i*, $\overline{\psi_n}$ is the true spatial average value of EOF mode *n*, *M* is the number of EOFs used, $\langle E_i^2 \rangle$ is the previously estimated mean data error variance for coarse grid box *i* and $\langle . \rangle$ represents an ensemble average. EOFs, being based on the spatial covariance structure of the data, are better for large-scale analyses than analytical correlation-versus-distance functions as these cannot take account of geographically-varying teleconnections. In both averaging steps, data weights, w_i, were optimised according to the data covariance structure, the spatial distribution of data and the error variances. Grid box error variances comprise representivity uncertainties (from *Jones et al.*, 1997) and measurement errors and are inversely related to the number of observations in each grid box. We estimated the two standard error (2σ) measurement error to be 0.4°C in any single daily LAT observation and used published error estimates (*Kent et al*., 1999) for SST observations. Correlation functions and EOFs were calculated from globally complete anomaly fields for 1948-1999. These were created using a Poisson technique (*Reynolds*, 1988) to interpolate HadCRUTv with 2m temperature anomalies from the National Centers for Environmental Prediction Reanalysis (*Kalnay et al*., 1996). This allowed truly global or hemispheric error estimates to be made. As shown by the equation, $\overline{\varepsilon^2}$ depends on the data errors and the difference between the true average of each EOF and the weighted average of its values where data exist. Thus the term in square brackets is smaller (greater) if the mode is well (poorly) sampled. The size of the eigenvalues also influences the error estimate. In a non-stationary climate, EOFs based on the relatively well observed 1948-1999 period may not be ranked similarly in earlier periods, particularly the first EOF which describes long-term warming. Thus we reordered our fixed EOF modes through time by projecting each of the 1948-1999 EOFs onto the annual HadCRUTv anomalies in overlapping 52-year periods: 1875-1926, 1876-1927, …, 1948-1999. The variances of each projection gave new eigenvalues for each period, reordering the EOFs. We truncated the reordered EOFs to retain approximately 90% of the 52-year variance, so that the number of EOFs retained varied in time. This was done because the highest order EOFs represent only noise and truncation gives a more stable estimate of covariance. We calculated the effect of the truncation on $\overline{\varepsilon^2}$ for the period 1948-1999 when complete data were available. The value of $\overline{\varepsilon^2}$ was reduced by 8%, so estimates of $\overline{\varepsilon^2}$ calculated from the above equation were increased accordingly in all years. The truncated set of EOFs was used to calculate the optimal average for the 26th year of each 52-year period, though the EOFs selected for 1948-1999 were used for all years from 1974 onwards, and the EOFs selected for 1875-1926 were used for years prior to 1900 because data were sparse. This procedure little affects the optimal averages, but realistically increases $\overline{\varepsilon^2}$ in the nineteenth century. Optimal methods can be biased towards zero anomaly and so underestimate climate change (*Hurrell and Trenberth*, 1999), but our 52-year periods are long enough to isolate a global warming EOF as an important source of variance.

The uncertainties given by RSOA due to data gaps and random errors (Figure 1a) were augmented using published estimates of global uncertainties associated with urbanization effects (e.g. *Jones et al*., 1990), with changes in the exposure of land thermometers (*Parker*, 1994), and with the SST bias-corrections (*Folland and Parker*, 1995). The urbanization uncertainty could be regarded as one sided: stations cannot be "too rural" but may inadvertently be "too urban" (*Jones et al*., 1990; *Peterson et al*., 1999). However, because some cold biases are also possible in adjusted semi-urban data, we conservatively model this uncertainty as symmetrical about the optimum average. We assume that the global average LAT uncertainty (2σ) owing to urbanisation linearly increases from zero in 1900 to 0.1°C in 1990 (*Jones et al*, 1990), a value we extrapolate to 0.12°C in 2000 (Figure 1a). We have not accounted for other changes in land use as their effects have not been assessed. Uncertainties in LAT due to non-standard thermometer exposures were estimated using published field experiments and the history of observing and screen usage (*Parker*, 1994). Our estimated 2σ errors take into consideration the range of local biases reported experimentally and the incompleteness of information on the exposures used operationally. Our estimated 2σ extratropical hemispheric uncertainties are 0.2°C before 1900, declining to zero by 1930 because universal standard exposures are assumed and random errors in standard exposures are already included in the optimal averaging. For the tropics, 2σ uncertainties are estimated as 0.4°C until 1930, declining linearly to zero by 1950. The uncertainties due to urbanisation and thermometer screens are mutually independent in a given year, so their variances were multiplied by the annually-varying land-based fraction of total data area in the region and added to the RSOA uncertainties. Similarly, the SST bias correction uncertainties reported by *Folland and Parker* (1995) were added after multiplication by the fraction of total data area covered by SST boxes.

Annual and decadal global surface temperature anomalies in Figures 1b and c are shown with 2σ confidence intervals. Uncertainties due to data gaps and random errors alone are often much less than those due to the other forms of uncertainty (Figure 1a), and considerably less than estimated previously using different methods (*Jones et al*., 1997). The more robust RSOA method used here minimises these errors. In particular, RSOA seeks to produce the most accurate estimate of the average by applying smallest weights to the most uncertain values, and is therefore expected to produce smaller errors than other methods. As a result, the global RSOA average may not be exactly equal to averages calculated by different methods, e.g. taking the average of the two hemispheres.

Tests of Bias Corrections to Sea Surface Temperature

The credibility of the optimum averages before 1942 depends considerably on the accuracy of the bias corrections to SST, even though we have incorporated the published uncertainties. The largest known source of bias occurred before late 1941 after which time SST began to be sampled mostly via ship engine intakes, common to this day, rather than using uninsulated or partly-insulated buckets (*Folland and Parker*, 1995). This change necessitated the development of geographically and seasonally complex corrections to pre-1942 SST data. Temporal changes in these bias corrections, both globally and locally, are only significant on decadal and longer time scales. The global annual mean bias correction increases steadily from 0.17°C in 1872 to 0.30°C in 1900 and

0.39°C in 1920, remaining around 0.4°C until 1941. Smaller, more recent, variations in measurement practice (*Kent et al*., 1999) are not assessed here.

The accuracy of the SST bias corrections has been largely confirmed near New Zealand and Japan (*Folland and Salinger*, 1995; *Hanawa et al*., 2000), but their global accuracy has not been assessed. We achieve this here through simulations of LAT using the HadAM3 atmospheric general circulation model (*Pope et al*., 1999). One ensemble of six experiments, 1870-1998, differing only in their initial atmospheric conditions, was forced with observed bias-corrected SST; another ensemble of four experiments, 1870-1945, was forced with uncorrected SST. Changes in greenhouse gases, ozone, anthropogenic aerosols or other natural forcings were not included in the simulations. Our analyses begin in 1872 to allow the model to adjust to the imposed SSTs and we use simulated global and large regional mean LAT anomalies before 1942 to test the corrections.

LAT anomalies were generally markedly lower when the model was forced with uncorrected, colder, SST than when it was forced with corrected SST (Figure 2). Corrected SSTs gave global average LAT anomalies much closer to those observed, especially between 1900 and 1941 (Figure 2a). Observed and ensemble mean simulated temperatures also agreed closely on hemispheric and large-regional scales (Figure 2b-e), especially on decadal timescales, when the model was forced with corrected SST, but not when forced with uncorrected SST. Internal variability in the model, although sometimes marked, did not obscure differences between the ensembles. Agreement in the tropics is noteworthy as the tropics have some of the largest annually averaged positive bias corrections to SST (*Folland and Parker*, 1995). However, observed tropical LAT anomalies are nearly 0.1°C warmer between 1900 and 1940 than LAT anomalies simulated with corrected SST. This may be due to the use of thatched shed screens, which resulted in relatively higher thermometer readings than those obtained using modern Stevenson type screens (*Parker*, 1994). European temperatures are well reproduced after 1895; earlier observations may be biased cold by thermometer exposures (*Parker*, 1994). Discrepancies between decadally smoothed simulated and observed global mean LATs after 1970 are due to a lack of explicit anthropogenic forcings in this model (*Sexton*, 2001).

Figure 2f shows an apparent failure of the corrected-SST simulations before about 1930 over Australia. Pre-1910 Australian data are probably erroneously warm relative to modern data owing to different thermometer exposures (*Nicholls et al*., 1996b) with errors that may be larger than reflected in our global estimates of uncertainties due to thermometer exposures. The cause of the discrepancy between 1910 and 1930 is unclear; it might be due to less accurate SST data south of Australia. Newly adjusted LAT data for three stations in southeast Australia are less warm in the late nineteenth century (not shown) so that the corrected-SST simulations for the average of these southeast Australian stations agree better. So our simulations can apparently identify regions in which LAT observations may be strongly biased.

Without bias corrections to SST, ensemble mean LATs are significantly too cold between 1872 and 1941 in all major regions except extratropical S. America (Figure 3). With bias corrections, however, only two regions have ensemble means significantly too cold (Figure 3). One is the tropics where the differences are quite small but model internal variability is low; the difference may reflect non-standard observing techniques mentioned above. The other is the extratropical Southern Hemisphere, reflecting the biased Australian observations.

Linear Temperature Trends Since 1861 and 1901

Finally we assess, linearly, overall warming and its uncertainty in global and hemispheric combined LAT and SST for 1861-2000 and for 1901-2000. We calculated separate RSOA uncertainties for the hemispheres, but used time series of the remaining global uncertainties weighted according to the land- and ocean-based data area fractions in each hemisphere. Table 1 shows overall warmings derived from linear trends in annual data, with twice their standard errors. Restricted maximum likelihood methods were used to produce unbiased trend estimates and their confidence intervals, taking into account the serial correlation structure of the data about the trend and the above error estimates (*Diggle et al*., 1999). The noise about the trend is assumed to be normally distributed with a covariance structure described in general by an AR(1) model and a set of independent random errors whose variance is estimated as that of the annual observational uncertainty. An AR(2) model was needed for Northern Hemisphere temperatures between 1861 and 2000 when all the uncertainties were included. Confidence intervals in deduced trends are insensitive to the inclusion or omission of the thermometer exposure uncertainty, showing that they are dominated by the correlation structure of the annual residuals about the trend.

Our estimated global warming trends since 1861 or 1901 are about four and three times larger than their 2σ uncertainties respectively. Total warming over the two periods is similar, with more warming in the Northern (> 0.6ºC) than the Southern Hemisphere (near 0.5ºC). Our global and hemispheric time series are close to those previously reported by IPCC (*Nicholls et al*., 1996a) using simpler methods, except for warmer temperatures in the Northern Hemisphere in the 1860s, but our uncertainty estimates are new.

Conclusions

We have provided the first quantitative estimates of uncertainty in global temperature change that take account of all known major factors. We have also provided an important validation of both decadal mean global LAT and bias-corrected SST anomalies, as the former can be fairly accurately derived from the latter. The results, excluding uncertainties due to thermometer exposures and the small eigenvector truncation factor, have been incorporated into the Third IPCC Assessment Report (*Folland et al*., 2001).

Acknowledgements. This work was supported by the UK Ministry of Defence and the UK Dept. of the Environment, Transport and the Regions. PDJ is supported by the US Dept. of Energy. Through the Met Office authors, the work is Crown Copyright. The authors thank R.E. Livezey and M.R. Allen for valuable discussions.

References

Diggle, P. J., K.-Y. Liang and S. J. Zeger, *Analysis of longitudinal data*, 253 pp., Clarendon Press, Oxford, 1999.

Folland, C. K. and D. E. Parker, Correction of instrumental biases in historical sea surface temperature data, *Q. J. R. Meteorol. Soc., 121*, 319-367, 1995.

Folland, C. K. and M. J. Salinger, Surface temperature trends in New Zealand and the surrounding ocean, 1871-1993, *Int. J. Climatol*., *15*, 1195-1218, 1995.

Folland, C.K., T.R. Karl, J.R. Christy, R.A. Clarke, G.V. Gruza, J. Jouzel, M.E. Mann, J. Oerlemans, M.J. Salinger and S.-W. Wang, Observed Climate Variability and Change. In: *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change,* ed. J. T. Houghton et al., Cambridge Univ. Press, 2001.

Hanawa, K., S. Yasunaka, T. Manabe and N. Iwasaka, Examination of correction to historical SST data using long-term coastal SST data taken around Japan, *J. Meteorol. Soc. Japan*, *78*, 187-195, 2000.

Hurrell, J. W. and K. E. Trenberth, Global sea surface temperature analyses: multiple problems and their implications for climate analysis, modeling and reanalysis, *Bull. Am. Meteorol. Soc.*, *80*, 2661-2678, 1999.

Jones, P. D., P. Ya. Groisman, M. Coughlan, N. Plummer, W.-C. Wang and T. R. Karl, Assessment of urbanization effects in time series of surface air temperature over land, *Nature*, *347*, 169-172, 1990.

Jones, P. D., T. J. Osborn, and K. R. Briffa, Estimating sampling errors in large-scale temperature averages, *J. Clim.*, *10*, 2548-2568, 1997.

Jones, P. D., T. J. Osborn, K. R. Briffa, C. K. Folland, E. B. Horton, L. V. Alexander, D. E. Parker and N. A. Rayner, Adjusting for sampling density in grid-box land and ocean surface temperature time series. *J. Geophys. Res.*, 106, 3371-3380, 2001.

Kalnay, E. and 21 co-authors, The NCEP/NCAR 40-year reanalysis project, *Bull. Amer. Meteorol. Soc. 77*, 437-471, 1996.

Kent, E. C., P. G. Challenor and P. K. Taylor, A statistical determination of the random observational errors present in Voluntary Observing Ships' meteorological reports, *J. Atmos. Oceanic Technol.*, *16*, 905-914, 1999.

Nicholls, N., G. V. Gruza, J. Jouzel, T. R. Karl, L. A. Ogallo and D. E. Parker, Observed Climate Variability and Change, in *The IPCC Second Scientific Assessment*, ed. J. T. Houghton et al., pp.133-192, Cambridge Univ. Press, New York, 1996a.

Nicholls, N., R. Tapp, K. Burrows and D. Richards, Historical thermometer exposures in Australia, *Int. J. Climatol.*, *16*, 705-710, 1996b.

Parker, D. E., Effects of changing exposure of thermometers at land stations, *Int. J. Climatol.*, *14*, 1-31, 1994.

Peterson, T. C., K. P. Gallo, J. Lawrimore, T. W. Owen, A. Huang and D. A. McKittrick, Global rural temperature trends, *Geophys. Res. Lett.*, *26*, 329-332, 1999.

Pope, V. D., M. L. Gallani, P. R. Rowntree and R. A. Stratton, The impact of new physical parametrizations in the Hadley Centre climate model HadAM3, *Clim. Dyn.*, *16*, 123-146, 1999.

Reynolds, R.W., A real-time global sea surface temperature analysis, *J. Clim.*, *1*, 75-86, 1988.

Sexton, D. M. H., Estimation of anthropogenic signals in an atmospheric climate model, using the General Linear Model, Ph.D. thesis, University of Reading, Reading, U.K., 2001.

Shen, S. S., T. M. Smith, C. F. Ropelewski and R. E. Livezey, An optimal regional averaging method with error estimates and a test using Tropical Pacific SST data, *J. Clim.*, *11*, 2340-2350, 1998.

Smith, T. M., R. W. Reynolds and C. F. Ropelewski, Optimal averaging of seasonal sea surface temperatures and associated confidence intervals (1860-1989), *J. Clim.*, *7*, 949-964, 1994.

C. Folland, N. Rayner, D. Parker, I. Macadam, D. Sexton, Hadley Centre for Climate Prediction and Research, Met Office, Bracknell, Berkshire, RG12 2SY, UK. (email: chris.folland@metoffice.com)

S. Brown, Department of the Environment, Transport and the Regions, London, UK

T. Smith, National Climatic Data Center, Asheville, NC.

S. Shen, Department of Mathematical Sciences, University of Alberta, Edmonton, Canada

P. Jones, Climatic Research Unit, University of East Anglia, Norwich, UK

R. Jones, Commonwealth Scientific and Industrial Research Organization, Aspendale, Australia

N. Nicholls, Bureau of Meteorology Research Centre, Melbourne, Australia

(Received January 18, 2001; revised March 30, 2001;
accepted April 12, 2001.)

Table 1. Surface temperature changes (ºC) and their 2σ uncertainties, 1861-2000 and 1901-2000.

	Period	Globe	N. Hem.	S. Hem.
All uncertainties included	1861-2000	0.61±0.16	0.64±0.26	0.51±0.14
	1901-2000	0.57±0.17	0.64±0.22	0.48±0.15
Land thermometer exposure uncertainty excluded	1861-2000	0.61±0.16	0.63±0.22	0.51±0.14
	1901-2000	0.57±0.17	0.64±0.22	0.48±0.15

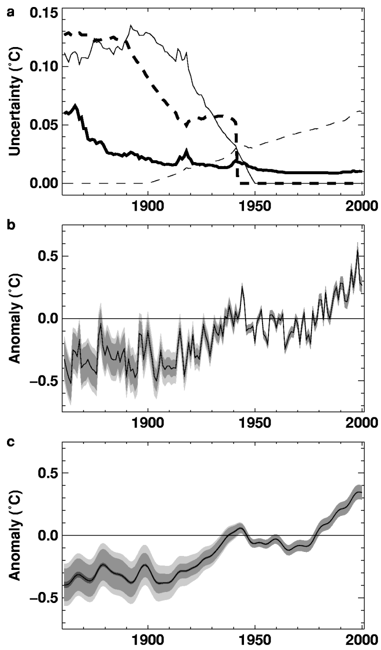

Figure 1. Global surface temperature anomalies, 1861-2000. **a** Uncertainties (2σ) due to: data gaps and random errors estimated by RSOA (heavy solid); SST bias-corrections (heavy dashes); urbanisation (light dashes); changes in thermometer exposures on LAT (light solid). LAT (SST) uncertainties are multiplied by the fraction of data area which is land (ocean) based. **b** Annual global surface temperature anomalies relative to 1961-1990 with 2σ confidence intervals and **c** decadal time series. Uncertainties in **b** are shown including (light shading) and excluding (darker shading) those due to changes in thermometer exposures. Averages in **c** were smoothed with a 21-term binomial filter and uncertainties estimated from annual values allowing for serial correlation. Dark shading: uncertainties from RSOA alone; intermediate shading: with added SST bias-correction and urbanisation uncertainties; light shading: all uncertainties.

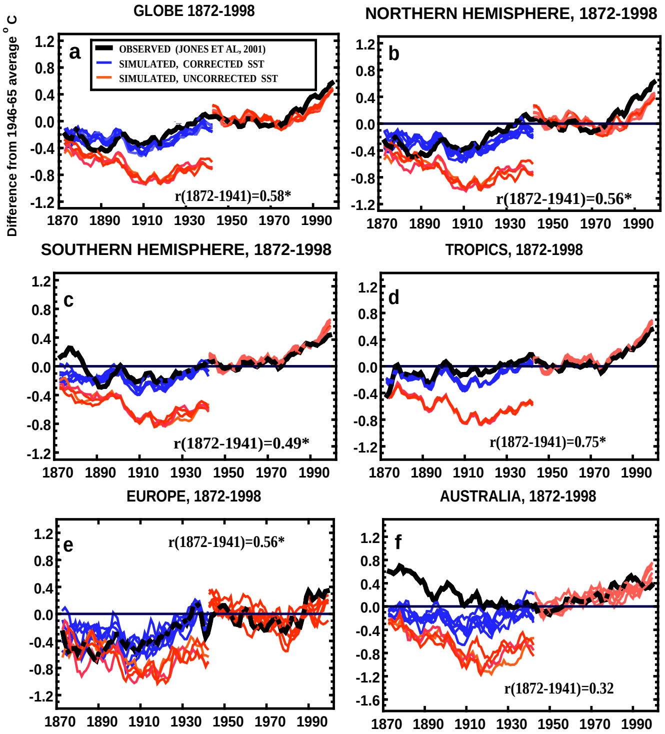

Figure 2. Simulated and observed global and large-region mean annual LAT anomalies (°C), 1872-1998, smoothed using a 21-year binomial filter. Modelled (observed) values are anomalies from ensemble mean (observed) averages for 1946-1965. Modelled values are sampled only where observations exist. Each ensemble member is shown. Smoothing on the simulations of 1872-1941 was carried out separately from that on the simulations of 1942-1998, to allow direct comparison between the simulations using corrected SST and those using uncorrected SST which ended in 1941; hence discontinuities in 1941/2. Correlations are of unsmoothed annual LAT observations with the unsmoothed annual ensemble mean simulated using corrected SST. Asterisk: significance at the 1% confidence level.

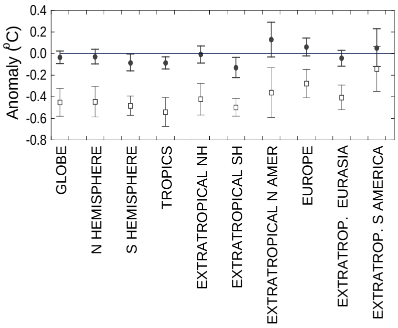

Figure 3. Mean regional differences between model ensemble means of annual LAT anomalies (°C) and observations for 1872-1941 with 2σ confidence intervals allowing for serial correlation. Open squares (closed circles): simulations with uncorrected (corrected) SST. Tropics 20°S-20°N; extratropical N. Hemisphere 20-90°N; extratropical S. Hemisphere 20-90°S; extratropical N. America 30-90°N, 170-60°W; Europe 30-90°N, 20W-30°E; extratropical Eurasia 30-90°N, 20W-180°E; extratropical S. America 30-90°S, 90-30°W.